Computational Models of Personality Recognition through Language

François Mairesse & Marilyn Walker
University of Sheffield, United Kingdom

- **Recognize personality**
  - **From written language**
  - **From conversations**

  ➢**Improve user modeling in computer systems**
    - Dialogue systems
    - Virtual agents
    - Intelligent tutoring systems

- **Most essential personality traits?**
- **Factor analysis of descriptors → 5 dimensions** (Norman, 1963)

  - **Extraversion**
    - Sociability, assertiveness vs. quietness
  - **Emotional stability**
    - Calmness vs. neuroticism, anxiety
  - **Agreeableness**
    - Kindness vs. unfriendliness
  - **Conscientiousness**
    - Need for achievement, organization vs. impulsiveness
  - **Openness to experience**
    - Imagination, insight vs. conventionality

Personality Correlates for Recognition

- **Attitude toward machines** (Sigurdsson, 1991)
  - E.g. neurotics have problems using computers

- **Academic motivation** (Komarraju & Karau, 2005)
  - Extravert and open students are more engaged in learning, conscientious achieve more
  - ➤ Training systems

- **Leadership** (Hogan et al., 1994)
  - High on extraversion, stability, conscientiousness and openness
  - ➤ Leader identification in meetings

- **Relationship success** (Donnellan et al., 2004)
  - E.g. both partners high on openness to experience
  - ➤ Partner matching in dating websites

Language and Personality

- **Linguistic markers of extraversion** (Furnham, 1990)
  - Talk more, faster, louder and more repetitively
  - Lower type/token ratio
  - More positive emotion words (Pennebaker & King, 1999)
    - E.g. *happy*, *pretty*, *good*

- **Emotional instability** (Pennebaker & King, 1999)
  - 1st person singular pronouns

- **Conscientiousness** (Pennebaker & King, 1999)
  - Fewer negations and negative emotion words

- **Low but significant correlations**

- ➤ **What about non-linear relations?**
- ➤ **No-one has tried to recognize personality on unseen subjects**

Data driven approach:

1. Collect individual corpora
2. Collect associated personality ratings
3. Extract features from the texts
4. Build statistical models of the personality ratings
5. Test the models on unseen individuals

Data driven approach:

1. **Collect individual corpora**
2. **Collect associated personality ratings**
3. **Extract features from the texts**
4. **Build statistical models of the personality ratings**
5. **Test the models on unseen individuals**

Corpus 1: Stream of Consciousness Essays
(Pennebaker & King, 1999)

- **2,479 essays over 7 years (1.9M words)**
- **Self-report personality assessment**
  - **Five Factor Inventory questionnaire** (John et al., 1991)

I see myself as someone who...
1. ...Is talkative
Disagree 1○ 2○ 3○ 4○ 5○ Agree

Introvert	Extravert
I've been waking up on time so far. What has it been, 5 days? Dear me, I'll never keep it up, being such not a morning person and all. But maybe I'll adjust, or not. [...]	I feel like I was born to do BIG things on this earth. But who knows... There is this Persian party today. My neck hurts. [...]

Corpus 2: Daily Conversation Extracts
(Mehl, Golsing & Pennebaker, in press)

- **96 participants recorded for 2 days, wearing an Electronically Activated Recorder (EAR)**
  - **Self-report personality ratings**
  - **Averaged personality ratings from 7 observers (r = 0.84, p < 0.01)**

Introvert 🔊	Extravert 🔊
- I don't know man, it is fine I was just saying I don't know. - I was just giving you a hard time, so. - I don't know. - I will go check my e-mail. - I said I will try to check my e-mail, ok.	- Oh, this has been happening to me a lot lately. Like my phone will ring. It won't say who it is. It just says call. And I answer and nobody will say anything. So I don't know who it is. - Okay. I don't really want any but a little salad.

- **Essays or conversations?**
- **Self reports or observer reports?**

Datasets	Self reports	Observer reports
Written language	Yes	?
Spoken language	Yes	Yes

Data driven approach:

1. **Collect individual corpora**
2. **Collect associated personality ratings**
3. **Extract features from the texts**
4. **Build statistical models of the personality ratings**
5. **Test the models on unseen individuals**

Automatic Feature Extraction

- **Utterance type (initiative)**
  - **Utterance tags based on parse tree**
    - Command, back-channel, question or assertion (Walker & Whittaker, 1990)

- **Content and syntax**
  - **LIWC categories** (Pennebaker & Francis, 2001)
    - **E.g. Positive emotion words, swear words, 1st person pronouns**
  - **MRC Psycholinguistic database** (Coltheart, 1981)
    - **E.g. Familiarity, age of acquisition, concreteness**

- **Prosody**
  - **Voice pitch, intensity and speech rate**

Data driven approach:

1. Collect individual corpora
2. Collect associated personality ratings
3. Extract features from the texts
4. Build statistical models of the personality ratings
5. Test the models on unseen individuals

Statistical Personality Modelling

- ## Regression problem?
  - E.g. extraversion = 4.3 on a 1-5 scale
  - Linear regression, regression trees

- ## Classification problem?
  - E.g. introvert vs. extravert
  - Decision tree, Naïve Bayes, Nearest Neighbour, SVM

➢ Depends on task and adaptation capabilities

- **Ranking problem?**
  - E.g. X is more extravert than Y

- **RankBoost** (Freund et al. 2003)
  - Non-linear model using boosting
  - Computes a ranking score for each instance
  - Minimizes the ranking error in the training data
   - percentage of misordered instance pairs

Extravert

A
B
C

Introvert

→ **Ranking model** →

B
A
C

▷ **33.3% ranking error**

Data driven approach:

1. Collect individual corpora
2. Collect associated personality ratings
3. Extract features from the texts
4. Build statistical models of the personality ratings
5. Test the models on unseen individuals

Regression Results - Essays

- **Baseline: average personality score**

- **Accuracy metric: improvement (%) over the baseline's absolute error**

- **10 fold cross validation**
  - 90% of the data for training / 10% for testing

- **Results with self-reports: Models outperform the baseline for all traits (p < 0.05)**

- **BUT very small improvement**
  - **Between 0.7% (Extraversion) and 6.2% (Openness)**

- ➢ **What if we model spoken language?**
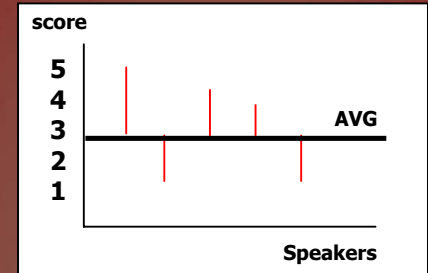
Regression Results - Conversation

- **Conversation data with self-reports**
  - **Never significantly outperform the baseline**

- **Conversation data with observer ratings**

	Improvement	Model
Extraversion	23.20%	M5' model tree
Emotional stability	3.92%	M5' regression tree
Agreeableness	None	
Conscientiousness	14.75%	M5' regression tree
Openness	None	
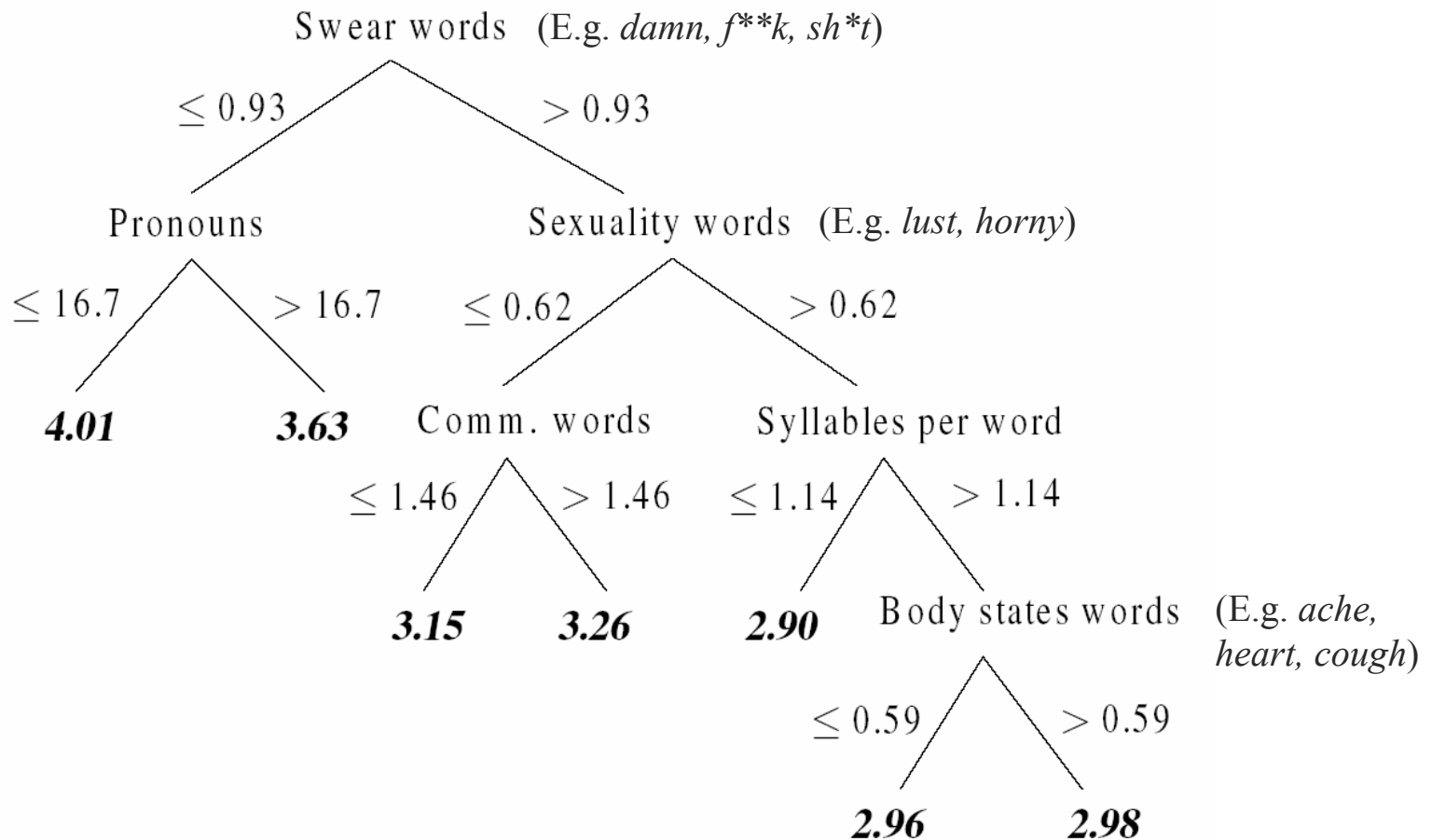
Binary Classification Results – Conversation
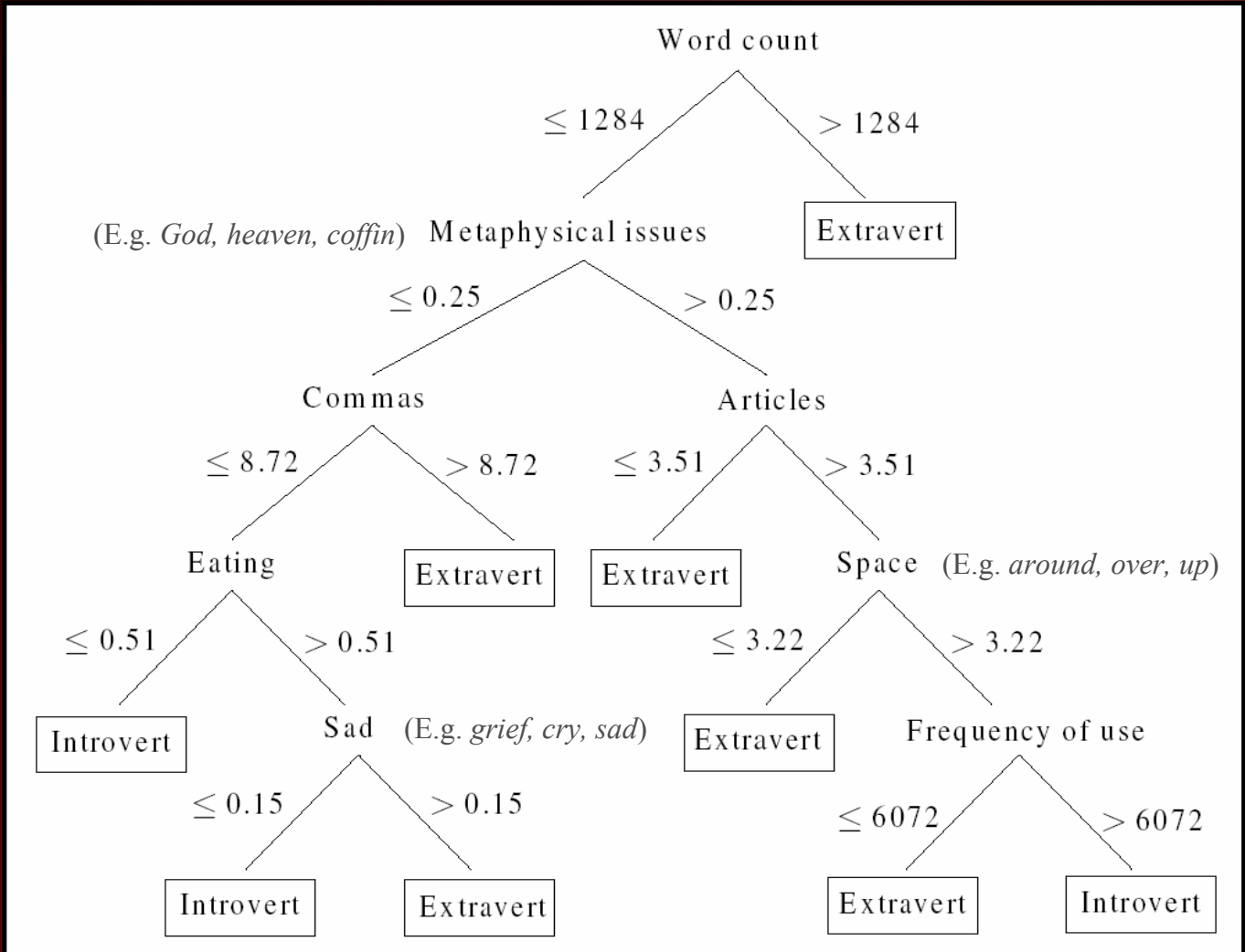
- **Observer reports**
- **Accuracy metric: correct classifications (%)**
- **Baseline: majority class (~ 50%)**
- **Naïve Bayes best model for all traits**

	Accuracy
Extraversion	**73.20**●
Emotional stability	**70.71**●
Agreeableness	55.08
Conscientiousness	**65.68**●
Openness	56.53

● significantly better than the baseline (two-tailed, $p < 0.05$)

- **67.26% accuracy**

- **Better than baseline (p < 0.05)**

Word count
- ≤ 1284 → Metaphysical issues (E.g. *God, heaven, coffin*)
- > 1284 → **Extravert**

Metaphysical issues
- ≤ 0.25 → Commas
- > 0.25 → Articles

Commas
- ≤ 8.72 → Eating
- > 8.72 → **Extravert**

Articles
- ≤ 3.51 → **Extravert**
- > 3.51 → Space (E.g. *around, over, up*)

Eating
- ≤ 0.51 → **Introvert**
- > 0.51 → Sad (E.g. *grief, cry, sad*)

Space
- ≤ 3.22 → **Extravert**
- > 3.22 → Frequency of use

Sad
- ≤ 0.15 → **Introvert**
- > 0.15 → **Extravert**

Frequency of use
- ≤ 6072 → **Extravert**
- > 6072 → **Introvert**

- **Baseline: random ranking (ranking error = 0.50)**
- **Paired t-test on a 10 fold cross-validation (two-tailed, $p < 0.05$)**
- **Self-reports models never outperform the baseline**

- **Observer models perform significantly better for all traits!**

	Ranking error	Feature set
Extraversion	0.26	Prosody
Emotional stability	0.39	MRC
Agreeableness	0.31	All
Conscientiousness	0.33	All
Openness	0.37	LIWC

- **Observed extraversion with prosodic features**
  - **Extraverts speak more, faster, with higher pitch**
  - **Introverts' voice pitch and intensity vary a lot**

Condition	α_i
Words-per-sec ≥ 0.73	1.43
Pitch-mean ≥ 194.6	0.41
Voiced-time ≥ 647.4	0.41
...	
Pitch-deviation ≥ 118.1	-0.15
Intensity-deviation ≥ 6.3	-0.18
Pitch-deviation ≥ 119.7	-0.47

Features of extraversion

Features of introversion

Sum → Extraversion ranking score

- **Observed conscientiousness with all features**
  - **Conscientious people**
    - Talk about their occupation (e.g. *work*, *class*, *boss*)
    - Use insight words (e.g. *think*, *know*, *consider*)
  - **Unconscientious people**
    - Swear a lot (e.g. *damn*, *f*ck*, *p*ss*)
    - Talk loud

Condition	α_i
Occupation ≥ 1.21	0.37
Insight ≥ 2.15	0.36
Positive feelings ≥ 0.30	0.30
Intensity-deviation ≥ 7.83	0.29
Num letters ≥ 3.29	0.27
…	
Swearing ≥ 0.93	-0.21
Swearing ≥ 0.17	-0.24
Religion ≥ 0.32	-0.27
Swearing ≥ 0.65	-0.31
Intensity-max ≥ 86.84	-0.50

- **Models performance better than baseline for extraversion, emotional stability, and conscientiousness**

- **Observed personality easier to model**
  - Self-reports are influenced by many factors, e.g. desirability of the trait

- **Spoken language with observer ratings produce best models**
  - Less constrained?

- **Regression results: (improvement over baseline)**

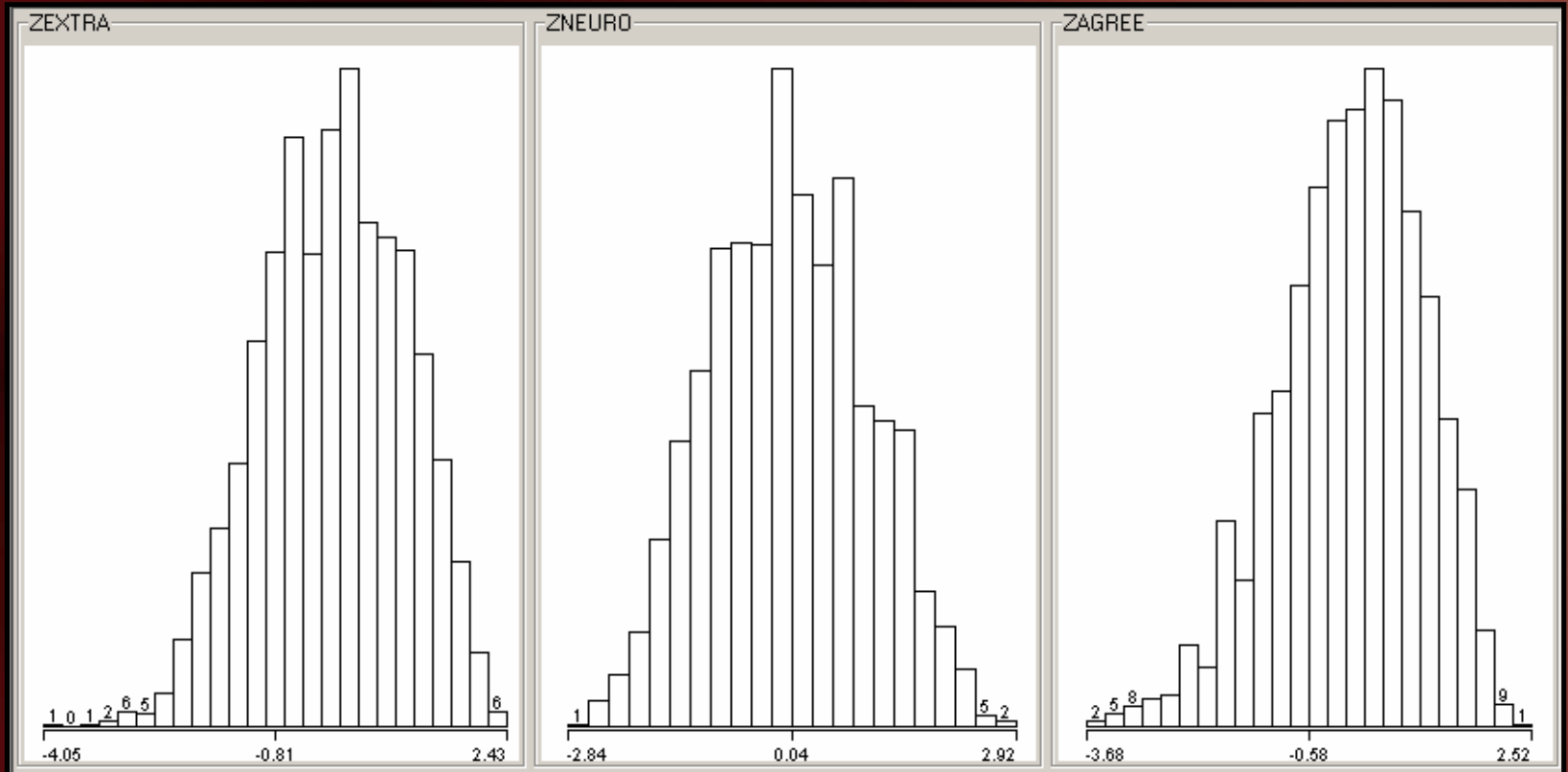| Datasets	Self reports	Observer reports
Written language	0.7% - 6.2%	?
Spoken language	N.S.	3.9% - 23.2%

References

References

- Sulloway F J 1999. Birth order. In: Runco M. A., Pritzker S. (eds.) Encyclopedia of Creativity 1: 189-202.
- Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change. *Journal of Personality and Social Psychology, 84,* 1041-1053.
- J. W. Pennebaker, L. E. Francis, and R. J. Booth, 2001. LIWC: Linguistic Inquiry and Word Count.
- W. T. Norman. 1963. Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality rating. J. of Abnormal and Social Psychology, 66:574–583.
- M. R. Mehl, S. D. Gosling, and J. W. Pennebaker. In press. Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. J. of Personality and Social Psychology.
- A. Furnham, 1990. Handbook of Language and Social Psychology, chapter Language and Personality. Winley.
- B. Donnellan, R. D. Conger, and C. M. Bryant. 2004. The Big Five and enduring marriages. J. of Research in Personality, 38:481–504.
- J.W. Pennebaker and L. A. King. 1999. Linguistic styles: Language use as an individual difference. J. of Personality and Social Psychology, 77:1296–1312.
- F. Heylighen and J.-M. Dewaele. 2002. Variation in the contextuality of language: an empirical measure. Context in Context, Special issue of Foundations of Science, 7:293–340.
- O. P. John and S. Srivastava. 1999. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin and O. P. John, editors, Handbook of personality theory and research. New York: Guilford Press.
- R. Hogan, G. J. Curphy, and J. Hogan. 1994. What we know about leadership: Effectiveness and personality. American Psychologist, 49(6):493–504.
- Y. Freund, R. Iyer, R. E. Schapire, and Y. Singer. 1998. An efficient boosting algorithm for combining preferences. In Proc. of the 15th ICML, p. 170–178.
- The interactive effects of conscientiousness and agreeableness on job performance. by Witt, L. A.; Burke, Lisa A.; Barrick, Murray A.; Mount, Michael K. from Journal of Applied Psychology. 2002 Feb Vol 87(1) 164-169.

- # Try the online demo!

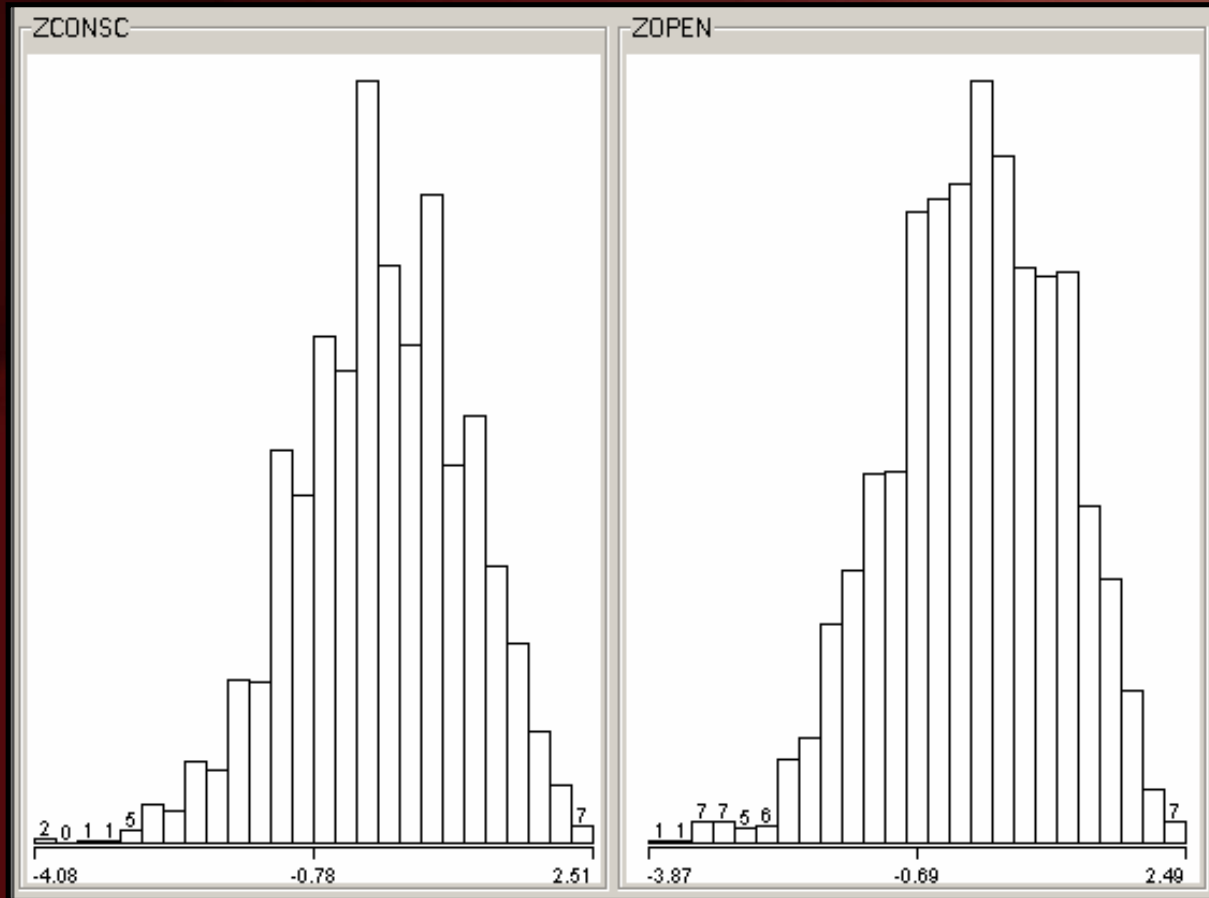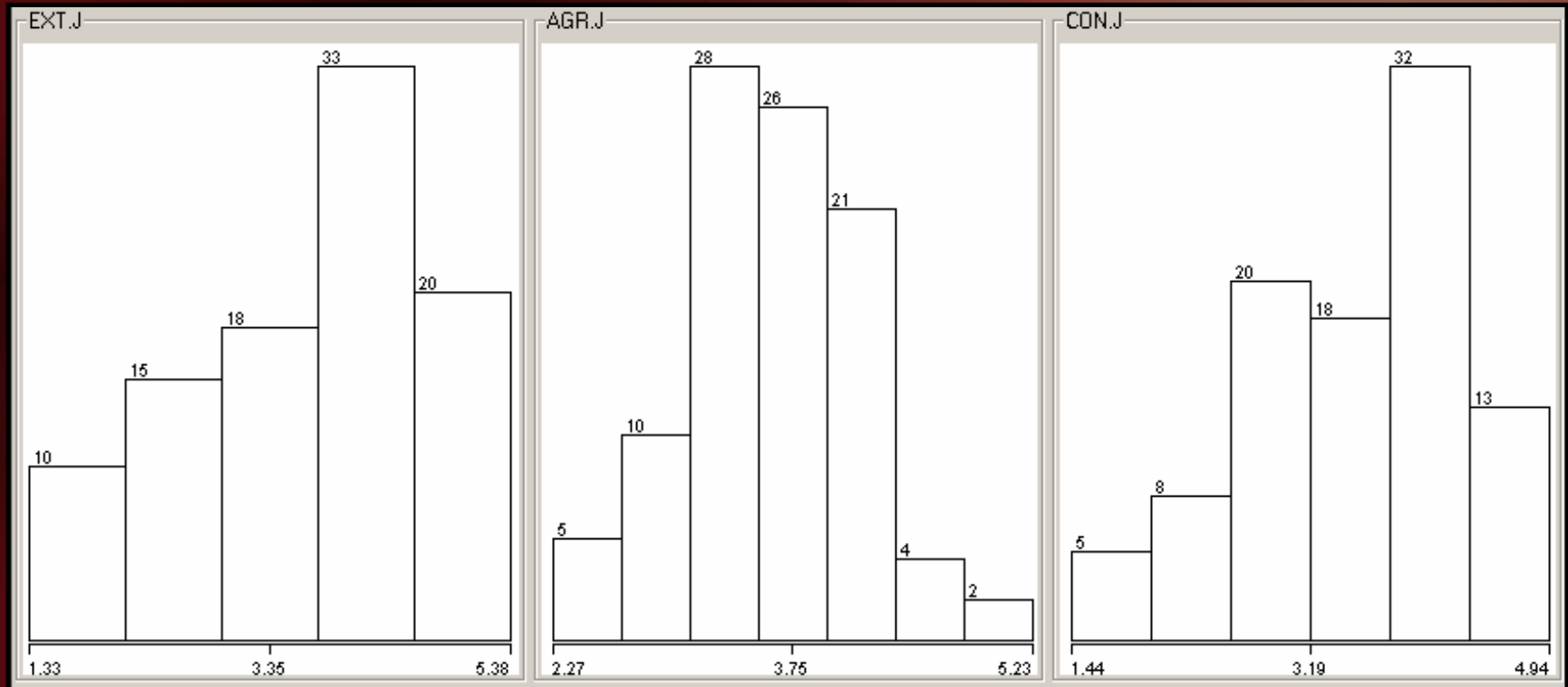http://www.dcs.shef.ac.uk/~francois/personality/demo.html

Thank you

Essays – Self Reports Distributions

Essays – Self Reports Distributions

EAR - Observer Ratings Distributions

- **Standard deviations between 0.5 and 1.0**